(Unconstrained)

Optimality : $\min_{x \in S} f(\underline{x})$

1st order: $f \in C^1$, $\bar{x}$ local min $\longrightarrow$ $\forall \underline{d} \in \mathbb{R}^n$ feasible : $\nabla^T f(\bar{x}) \underline{d} \geq 0$

2nd order: $f \in C^2$, $\bar{x}$ local min $\longrightarrow$ $\begin{cases} \nabla^T f(\bar{x}) \underline{d} \geq 0 & \underline{d} \text{ feasible direction} \\ \nabla^T f(\bar{x}) \underline{d} = 0 \Rightarrow & \underline{d}^T \nabla^2 f(\bar{x}) \underline{d} \geq 0 \end{cases}$

Suff. : $f \in C^2$, $\bar{x} \in \text{Int}(S)$: $\nabla f(\bar{x}) = \underline{0}$, $\nabla^2 f(\bar{x})$ p.d. $\longrightarrow$ $\bar{x}$ (strict) local min

Nec. : CONVEX PB.

$f \in C^1$ convex, $\underline{x}^*$ global min $\Longleftrightarrow$ $\nabla^T f(\underline{x}^*)(\underline{y} - \underline{x}^*) \geq 0 \quad \forall \underline{y} \in C$

Methods :

1. line search methods
2. Gradient method
3. Newton method
4. Conjugate direction method
5. Quasi-Newton method : DFP
6. Quasi-Newton method : BFGS

1.
- search directions : $\underline{d}_k$
- step length : α_k
  (Wolfe + Bisection)
- convergence

2. slow convergence

$\underline{d}_k = -\nabla f(\underline{x}_k)$

$\alpha_k = \dfrac{\underline{d}_k^T \underline{d}_k}{\underline{d}_k^T Q \underline{d}_k^T}$

$f(\underline{x}) = \frac{1}{2} \underline{x}^T Q \underline{x} - \underline{b}^T \underline{x}$

\+ convergence
\+ extension to arb. funct.s

3. heavy computation

$\underline{d}_k = -H^{-1}(\underline{x}_k) \nabla f(\underline{x}_k)$

$\alpha_k = 1$

$H(\underline{x}) = \nabla^2 f(\underline{x})$

Alternative interpr.

$\min f(x) \Longleftrightarrow f'(x) = 0$
method of the tangents

\+ convergence
\+ extensions

4. $q(\underline{x}) = \frac{1}{2} \underline{x}^T Q \underline{x} - \underline{b}^T \underline{x}$

$\alpha_k = -\dfrac{g_k^T d_k}{\underline{d}_k^T Q \underline{d}_k}$, $g_k = Q \underline{x}_k - \underline{b}$

$\underline{d}_k \in \{\underline{d}_i\}_{i=0}^{n-1}$ non-zero mutually Q-conjugate directions

$\underline{x}_n$ = global optimum of $q(\underline{x})$

\+ conjugate GRADIENT mth. for $\{\underline{d}_i\}_{i=0}^{n-1}$

$\begin{cases} \underline{d}_{k+1} = -g_{k+1} + \beta_k \underline{d}_k \\ \beta_k = \dfrac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \end{cases}$

\+ arbitrary functions (PR > FR)
\+ convergence
\+ preconditionated conj. gradient mth.

BROYDEN FAMILY

5. $H_{k+1} \underline{\gamma}_k = \underline{\delta}_k$ secant cond.

$\approx [\nabla^2 f(\underline{x}_k)]^{-1}$

- rank 1 update formula
- rank 2 update formula
  (+ curvature condition)

\+ properties & links

6. $B_{k+1} \underline{\delta}_k = \underline{\gamma}_k$

$\approx \nabla^2 f(\underline{x})$

- "rank 2 updated formula" for B_{k+1}
- BFGS for H_{k+1}

\+ properties & links